DOCUMENT RESUME

ED 202 918 TM 810 450

AUTHOR Horne, Leslie V.; Garty, Mary K.
TITLE What the Test Score Really Reflects: Observations of Teacher Behavior During Standardized Achievement Test Administration.
SPONS AGENCY Carnegie Corp. of New York, N.Y.
PUB DATE 81
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; Elementary Education; Elementary School Teachers; *Standardized Tests; *Teacher Behavior; *Teacher Influence; *Testing; Testing Programs; *Test Validity

ABSTRACT

This study focuses on standardized test administration in five elementary schools. The authors discuss how teachers affect test validity by altering standard test procedures in both planned and incidental ways. Each of the 16 voluntary teachers who participated in this study was observed in her/his classroom administering certain subtests of the Metropolitan Achievement Tests or the Stanford Achievement Test. Information gained from interviews with the teachers was used along with the observational data in order to better understand the teachers' behavior. Although the observations were begun expecting teachers to adhere strictly to the directions as outlined in the test manual, a wide range of teacher behavior was noted across the sixteen classrooms--nonstandard teacher behavior being observed in over one-half of the classrooms. These observations fell into three groups: (1) teachers who followed the directions closely; (2) teachers whose deviations were not systematic and appeared to be unintentional; and (3) teachers who consciously and purposely varied the test administration procedures. Findings highlight the importance of further study of teacher-introduced factors that may affect test validity. (Author/RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED202918

WHAT THE TEST SCORE REALLY REFLECTS: OBSERVATIONS OF TEACHER BEHAVIOR DURING STANDARDIZED ACHIEVEMENT TEST ADMINISTRATION

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Leslie V. Horne and Mary K. Garty

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. V. Horne & M. K. Garty

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Learning Research and Development Center

University of Pittsburgh

1981

TM 810 454

Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981. The research reported herein was supported by the Carnegie Corporation of New York. The opinions expressed do not necessarily reflect the position or policy of the Carnegie Corporation, and no official endorsement should be inferred.

Introduction

Standardized achievement testing continues to be an issue of concern in educational research circles. Most studies in this area focus on the uses and abuses of test results. In this study, however, the testing situation itself will be examined to discover how closely teachers adhere to testing procedures as outlined in the manuals which accompany these tests. Our interest in test administration stems from our conviction that variations in procedures, of course, affect test results. What became more interesting was to discover why those teachers who deviate from the norm do so.

Background Characteristics

The authors observed standardized test administrations in sixteen classrooms in five elementary schools. Two of the schools are in a small suburban school district and three schools are in a large urban district. The communities served by the schools range from upper-middle to lower socioeconomic status as measured by free lunch eligibility figures and the mean home values for the area surrounding the schools (obtained from a 1979 local realty company survey). Within the suburban schools, the student population is 98% white. Among the urban schools, one is 96% black, one 91% white, and one 94% white. Achievement test scores of the two suburban schools and one urban school are above national norms; two urban schools score below national norms (see Table 1).

Method and Procedures

Each teacher who participated in this study was observed in her/his classroom administering certain subtests of the Metropolitan Achievement Tests or the Stanford Achievement Test. It is important to note that the teachers agreed voluntarily to be observed. Several teachers refused to participate, citing the presence of an observer as distracting to their students. Average length of observation was 30 minutes. Notes on the observations were written immediately after the observer left the classrooms.

The authors do not claim to have randomly sampled the population of teachers in these schools. Nevertheless, we have observed at least one teacher at each elementary grade level. Participating teachers also differed by years of experience, race and sex (see Table 1).

The present study is part of a larger ongoing investigation of the uses of standardized tests in schools[1]. Data for the background study were obtained through two open-ended interviews with each participating teacher. The sample for this larger study included 146 teachers. Information gained from these interviews will be used along with the observational data in order to better understand the teachers' behavior.

---

[1] A summary of the results of the background study can be found in the May, 1981 Phi Delta Kappan (in press).

Results

The authors began these observations expecting teachers to adhere strictly to the directions as outlined in the test manual. In fact, a wide range of teacher behavior was noted across the sixteen classrooms.

The observations fell into three groups. The six Group I teachers closely followed the directions that accompany each test. Group II is composed of three teachers whose deviations were not systematic and appeared to be unintentional. For example, one teacher left the room for a short time. Another teacher inadvertently hampered her students' progress by instructing them to raise their hands if they had a question or needed a new pencil, then failed to place herself in a position where she could readily see the raised hands. Some children lost up to a half minute of work time waiting to get the teacher's attention.

Group III teachers consciously and purposely varied the test administration procedures. These seven teachers will be discussed at some length as they are the focus of our concern and their actions must be described in detail to be appreciated.

Teacher A. (School 3, Grade 5, Math Concepts and Computation): This teacher read the directions verbatim from the teachers' manual, but three unusual actions took place in his classroom while the standardized test was being administered. First,

Teacher A sent a child, who had apparently finished the test early, out of the classroom on an errand. Shortly after that, and despite the "TESTING. DO NOT DISTURB" sign on the door, the principal came into the room. Teacher A and the principal carried on a short conversation in normal tones of voice at the front of the classroom. While another section of the test was being given, several children who finished the test early were allowed to carry on a whispering, gesturing conversation.

Teacher B. (School 4, Fourth grade, Math Computation): Teacher B read the instructions from the manual as directed except for the part regarding the DK (don't know) response. Instead of explaining that this option was to be used when students don't know the answer, the teacher told the class not to use the DK option. About 10 minutes into the testing session, the teacher walked around the class checking his students' work. Then he informed the students that they were confusing addition and multiplication signs and they should go back and check their answers. One child was singled out and told to go back and start the test over. Teacher B permitted students who had completed the test early to carry on conversations with each other. This teacher sat next to and attempted to carry on a conversation with the research observer while the students were taking the test.

Teacher C (School 3, Fourth grade, Math Concepts, Math Problem Solving): Teacher C, like Teacher B, read the directions from the Teachers' Manual except for the instructions concerning the DK option. His directions were: don't use the DK options. If you mark DK they're going to mark it wrong, so you might as well guess. I don't want you to use DK. The teacher again told the class to omit DK at the beginning of the second subtest.

Teacher D (School 1, First grade, Vocabulary): The vocabulary test is administered as follows: the teacher reads incomplete sentences to the children and then reads three words, one of which would complete the sentence correctly. The three words are printed in the children's test booklets and they are to blacken the circle above the correct word. On two occasions, when questioned by a student, Teacher D defined or gave synonyms for the words on the test. When asked by one child about the meaning of "chisel," the teacher not only defined the word, but pantomimed a hammer hitting something to clarify the meaning. This teacher also stopped in the middle of the vocabulary test to give the children an exercise break. While the manual suggests an exercise break between tests, one halfway through the test is, at least, unorthodox.

Teacher E (School 1, First grade, Reading). In the observed section of the test, the format is as follows:

1.	moon	grass	flower
2.	forest	daisy	dime
3.	pole	past	petals [2]

The students look at the picture, then blacken the space under the one word in each line that tells something about the picture. Teacher E read the directions from the manual and the children began the test. The students did not have erasers and were told to raise their hands if they needed to erase. The teacher then walked around the room stopping when children raised their hands to indicate they had questions or needed an eraser. Teacher E stopped by a child's desk and asked what was to be erased. The child apparently indicated a correct answer. The teacher said, "You want me to erase this?" Then, pointing to alternative answers she said, "Well look at these other words. Does this word tell you about the picture? Does this word tell you about the picture?" Despite the teacher's hints, the child insisted that the right answer was the one to be erased, and the teacher finally did erase it. Teacher E gave this special help to several other children during the test. The other unconventional aid this teacher gave was to go over the children's tests when they were finished, point out lines that were blank and tell the children they

---

[2] Due to the copyright restrictions on standardized test materials, the authors have constructed a facsimile item.

were to have an answer on each line.

Teacher F (School 3, Fourth grade, Math Computation, Math Concepts): This teacher did not read the directions from the manual, but explained the procedures in her own words. The children were told to guess rather than use the DK response and to make very sure the answer was not given before marking the NG option. The teacher walked around the room as the children worked, stopping occasionally to look over their tests. After stopping at one child's desk, the teacher told the student, "check number 9." The teacher then continued to look over the child's booklet and finally said, "Just start all over again." Teacher F stopped at several other students' desks, looked over their tests and said, "You're not being careful."

Teacher G (School 3, Third grade, Math Concepts): According to the manual, the teacher should have read the directions and then allowed the children to complete the test independently. Instead, this teacher disregarded the standard directions and introduced the test in the following way. Teacher G first told the class that this is not a reading test but a math test, and what is important is to learn what math concepts they know or don't know and not how well they can read the questions. The teacher then explained that for these reasons, she would go through the test and read all the questions to the class to be sure they didn't get any wrong because they couldn't read the question. The children were told not to begin writing answers until all the questions had been read to them. The teacher

then read all 45 questions to the children, going into detailed explanations concerning how to use accompanying figures or drawings to solve the problems.

Teacher G also told the children to skip certain problems because they "hadn't had that yet." One child questioned the teacher, saying she felt sure she could do a problem the teacher had told the class to skip. The teacher replied, "Well, try it if you think you can do it." Then after examining the problem again, the teacher said, "Oh, that's not so hard, you all can probably do it." After reading all the problems to the class, the teacher read the directions from the manual aloud and the students began to work.

It is important to recall that these observations were not done surreptitiously and that teachers involved were aware of our interests in standardized test use. Two teachers in Group III felt the need to explain their alterations and deviations to the observers. Essentially, however, what was observed in Group III teachers was an open, intentional and unself-conscious alteration of standardized test administration procedures and conventions. The following sections will be an analysis and discussion of the teachers' behavior in the testing situation.

Analysis

In this section, the authors will identify variables that discriminate among the groups of teachers, and then we will examine those characteristics shared across groups. Finally, we will focus on the behavior of the Group III teachers.

Group I teachers are all from middle to upper-middle class white schools that score above national norms on standardized achievement tests. Two teachers teach in School 2 in the suburban disrict; four teach in the urban School 5 (see Table 1).

The three teachers in Group II are from suburban School 1 and urban School 4. Both of these schools serve low-middle to low income communities. Students in the urban school score below national norms on standardized achievement tests. Both schools serve predominately white student populations. The suburban school's students score above national norms but their scores are not as high as those of the other elementary school in that district.

Four of the Group III teachers are from the low-achieving predominantly black urban school, one from the predominantly white low-achieving urban school and two from the lower income suburban School 1. While the students in School 1 score above national norms on standardized achievement tests, it is clear from our interview data that the teachers are aware that their school is the lower achieving elementary school in the district. Teachers in Schools 1, 3 and 4 are, therefore, alike in that they serve low-income populations and in their perceptions of the achievement level of their students vis-a-vis the other schools in their districts.

Data has been collected on a number of individual teacher variables. It is interesting to note that our sample is comprised of experienced teachers: fourteen of the sixteen had taught for seven or more years. Therefore, the length of teaching experience did not allow us to discriminate among our teacher groups, nor did race, sex,

or grade level taught.

It would be easy to conclude this discussion by criticizing the non-conforming teachers and accusing them of altering the testing situation in an attempt to make themselves look better. After all, it is the teachers of the low achievers who varied test procedures. We, however, feel very strongly that these teachers' motives were not selfish and not centered on teacher accountability. Our contentions are supported in two ways. First from our knowledge of the districts involved, gained through the teacher interviews and interviews with administrators, we know that, in fact, standardized test results are not used for teacher accountability and second, as mentioned earlier, these teachers are experienced and tenured teachers and relatively secure in their positions.

Discussion of the interview data will help to provide an alternative explanation of the Group 3 teachers' behavior.

The teachers who gave the most help, and/or otherwise changed the test directions, were Teachers C, D, E, F and G. All suggested in the background interviews that their students, for a variety of reasons, were placed at a disadvantage in the standardized achievement test setting. For example, in his interview, Teacher C stated that his students were not "test-wise." He said specifically that he tells his students not to use the DK option -- because the answer will be marked wrong anyway, so why not guess? And in fact he did just that when he was observed giving the test. Teacher C considers this deviation from test directions to be a way of "teach(ing) the children how to take a test." From his interview, it is clear that C sees it as his

responsibility to educate his students on good test taking techniques which includes guessing.

In her interview, Teacher G stated that she resented how her (urban) school district used the test results to make comparisons among schools that had different "environments." Teacher F expresses a concern when interviewed that her students don't concentrate as well as she would like and are careless workers. It is consistent then that she would exhort her students to "be careful" and that she would tell them to "just start over again" as she noticed errors in their test performance.

At the close of the testing session, Teacher E told the observer that she "hated the test because she knew the children could do well but they got nervous and didn't." During the test, this teacher gave her students additional aid in the form of hints about which answers to erase or not erase. (The teacher had the only eraser.)

Though Teacher B admonished his students against using "DK" and about confusing signs, both Teachers A and B conveyed an almost casual attitude about test administration. Each permitted students to carry on conversations in the classroom and carried on their own conversation with visitors during the testing session. In his interview Teacher A, who teaches at the predominantly black urban school, said that he saw the test as "not relevant for black students." He said, too, that he really didn't like the test but tries hard not to convey this attitude to his students. The behavior that we observed in the testing situation, coupled with the fact that he allowed students to make up parts of another achievement test while

his background interview was being conducted, nevertheless made his attitude about standardized testing clear to us.

Discussion

To summarize our findings about Group III: all Group III teachers significantly altered the standardized achievement test setting. It appears from our analysis, however, that the teachers' motivations varied. Five teachers (C, D, E, F and G) altered testing procedures in order to enhance their students' performance. The inescapable irony is that despite their efforts, students in two of the schools where unstandardized behavior was observed still score below national norms. The two remaining teachers expressed a certain disdain for standardized tests which is evident in their behavior.

These teachers by and large attempt to "run interference" for the deficiencies in either the test or in their students. They are concerned that the constraints of the standardized test situation will not permit the true achievement of their students to be demonstrated. That is, these teachers believe that their assistance results in, not inflated, but more accurate scores for their students. This altruistic motive is far removed from a preoccupation with teacher accountability.

Teachers are placed in an unusual circumstance when administering a standardized test. The are compelled to administer a test that may not match what they have taught, and for which they are to give only minimal help. It may be, in fact, that teachers who follow the rules and teachers who don't are just continuing a pattern of behavior they

would exhibit for any test they administer, even their own. The authors do have a sense, for example, that Group I teachers are, as one teacher put it, "do it as you're told teachers." These teachers would be reluctant to give extra help, hints or whatever no matter what the test situation. Though we did not have the opportunity to observe Group III teachers giving their own tests perhaps they would, again, steadfastly act to facilitate their students' test taking experiences. It must be very difficult for teachers to change a year-long pattern of behavior for the one or two occasions a year that they must administer standardized achievement tests.

Conclusion

Upon discovering that over one-half of our sample did not adhere to standardized testing procedures, the authors attempted through use of the background interviews as well as socioeconomic and achievement data on the schools, to explain teacher motivations. Our findings highlight the importance of further study of teacher-introduced factors that may affect test validity.

Table 1

Description of Schools and Teachers

	District	Teachers Observed	Grade Levels Observed	SES of School Community	Past Achievement Test Results
School 1	Suburban	4 teachers *--white women	1st and 4th grades	Low to lower middle	Above National Norms
School 2	Suburban	2 teachers --white women	Kindergarten	Middle to upper middle	Above National Norms
School 3	Urban	4 teachers *--1 black woman *--1 black man *--1 white woman *--1 white man	3rd, 4th, and 5th grades	Low	Below National Norms
School 4	Urban	2 teachers *--1 white man --1 white woman	2nd and 4th grades	Low	Below National Norms
School 5	Urban	4 teachers --3 white women --1 white man	1st, 4th, and 5th grades	Middle to upper middle	Above National Norms

*Group 3 teachers.